UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/713,576	11/12/2003	Shyam Kapur	017887-011010US	8249

20350 7590 08/26/2004

TOWNSEND AND TOWNSEND AND CREW, LLP
TWO EMBARCADERO CENTER
EIGHTH FLOOR
SAN FRANCISCO, CA 94111-3834

EXAMINER	
LE, DEBBIE M	
ART UNIT	**PAPER NUMBER**
2177	

DATE MAILED: 08/26/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/713,576	KAPUR ET AL.	
	Examiner	Art Unit	
	DEBBIE M LE	2177	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *12 November 2003*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-31* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-31* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *12 November 2003* is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☒ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *11/12/03*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Priority

Applicant's claim for domestic priority from U.S Patent Application Serial No. 60/460,222 which was filed on April 4, 2003, under 35 U.S.C. 119(e) is acknowledged.

Information Disclosure Statement

The information disclosure statement (IDS) submitted on 11/12/03 has considered by the examiner and has been placed in the application file. Please see attached PTO-1449.

Drawings

The drawings 2-4 are objected to because they fail to show necessary textual labels of features or symbols as described in the specification. For example, placing a label, "Server System", with elements 160 of Fig. 2, or "Search Module", with element 126 of Fig. 2, would give the viewer necessary detail to fully understand this element at a glance. A ***descriptive*** textual label for ***each numbered element*** in these figures would be needed to fully and better understand these figures without substantial analysis of the detailed specification. Any structural detail that is of sufficient importance to be described should be shown in the drawing. Optionally, applicant may

---

The italicized words from the prior art are mapped to the claim limitations.

wish to include a table next to the present figure to fulfill this requirement. See 37 CFR 1.83. 37 CFR 1.84(n)(o) is recited below:

"(n) Symbols. Graphical drawing symbols may be used for conventional elements when appropriate. The elements for which such symbols and labeled representations are used must be adequately identified in the specification. Known devices should be illustrated by symbols which have a universally recognized conventional meaning and are generally accepted in the art. Other symbols which are not universally recognized may be used, subject to approval by the Office, if they are not likely to be confused with existing conventional symbols, and if they are readily identifiable.

(o) Legends. Suitable descriptive legends may be used, or may be required by the Examiner, where necessary for understanding of the drawing, subject to approval by the Office. They should contain as few words as possible."

Specification

On page 7, lines 34, Applicants are requested to update the application serial No. "10/_____," to --10/712,307--.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

The italicized words from the prior art are mapped to the claim limitations.

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

Claims 1, 6-19, 28-31 are rejected under 35 U.S.C. 103(a) as being unpatentable over Steinkraus (US Patent 6,363,373 B1) in view of Bond et al (US Patent 6,539,348 B1).

As per claim 1, Steinkraus discloses the recited limitations as follows:

'A computer-implemented method of generating concept units from user search queries' *as the search queries input by the user wherein the search terms are broken into word tokens* [See col. 2, lines 40-45] and *the search engine uses the word tokens to generate a list of concept tokens* [See col. 2, line 34]. The list of content tokens is used to *search for concepts* [See col. 2, line 58-59], the method comprising:

The italicized words from the prior art are mapped to the claim limitations.

'receiving a plurality of queries' *as preprocessing search queries input by the users* [See col. 7, line 30 and line 36], 'each query comprising a string of one or more words' *as the query contains words, word-like strings, numbers* [See col. 7, lines 38-39];

'tokenizing each query string to produce one or more tokens for each query' *as the query is read and broken into word tokens* [See abstract, lines 13-14], 'wherein said tokens for said queries form an initial set of units' *as the queries are broken and produced a series of word tokens* [See col. 5, lines 35-36];

'until a convergence condition is satisfied' *as the concept tokens weights are less than a threshold value* [See col. 8, lines 8-10].

Steinkraus does not explicitly teach combining units from the initial set of units that appear adjacent each other in a query to form a second set of units, validating the second set of units, repeating the steps of combining and validating one or more times using the second set of units in place of the initial set of units, wherein a final set of units is formed once the convergence condition has been satisfied and storing the final set of units to a memory.

However, Bond teaches, an analogous system for *tokenizing words of a natural language* [See abstract, lines 2-3]. In particular, Bond teaches:

'combining units from the initial set of units that appear adjacent each other in a query to form a second set of units' *as a token merging program for merging the sequence of a phrase or clause* [See col. 5, line 51, line 35, col. 1, line 26], *the processor iterates matching consecutive tokens against the first set of rule* [See Fig. 1, # 26, col. 5, lines 26-27]. *The processor will continues to iterate through the tokenized*

The italicized words from the prior art are mapped to the claim limitations.

of words until further word tokens can be made by processor [See col. 5, lines 27-31]. Bond explained *as the processor performs matching consecutive tokens in the resulting narrower set of possible syntactic interpretations against a second set of rules* [See col. 5, lines 32-33];

'validating the second set of units' *as the processor again, processes the second set of rules until no further narrowing of possible syntactic interpretations* [See col. 5, lines 36-37];

'repeating the steps of combining and validating one or more times using the second set of units in place of the initial set of units' *as the process proceeds into reiteration program code* [See col. 5, lines 38-40], 'wherein a final set of units is formed once the convergence condition has been satisfied' *as the completion of the iterative inductive processing output the syntactic possible interpretations marked text* [See Fig. 1, # 38, col. 5, lines 42-44]; and

'storing the final set of units to a memory' *as the syntactically marked text is output and retained in a software object to permit further analysis* by the users [See col. 6, lines 13-17].

Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to combine the teachings of the cited references to implement the steps of combining units, validating all units that adjacent each other, and repeating the steps of combining until a convergence condition is satisfied, then storing the final set of units to memory. Bond's teachings would allow users of Steinkraus's system to readily determine the appropriate syntax for a natural language query to obtain meaningful

---

The italicized words from the prior art are mapped to the claim limitations.

interpretations and analyses of enormous amount of information to conveniently search for the concepts associated with any given user's needs as suggested by Bond [See col. 1, lines 29-40, col. 2, line 62-63].

As per claim 6, Bond teaches 'including generating unit extensions using the final set of units' *as the completion of syntactic possibilities interpretations are passed on to deductive token merging code. The deductive code reviews possible interpretations and determines which ones are remaining for the inductive merging processes* [See col. 5, lines 59-64].

As per claim 7, Bond teaches 'wherein generating unit extensions includes identifying units that are subsets of other units' *as all words in a sentence are tokenized into multiple words* [See col. 4, lines 20-21]. Because Bond teaches that each of word in a sentence is a '*word token*', which is equivalent to the claim limitation "unit". Bond disclosed that an article: "*United States*" [See col. 4,line 19] is tokenized into word tokens, "*United*" and "*States*". Therefore, the "*United*" is a subset of other "unit", wherein the other unit is *a "States".*

As per claim 8, Bond teaches 'storing the unit extensions to the memory' *as the syntactically marked text is output and retained in a software object to permit further analysis* [See col. 6, lines 13-15].

The italicized words from the prior art are mapped to the claim limitations.

As per claim 9, Bond teaches 'generating unit associations using the final set of units' *as a listing of words and their associated part of speech is analyzed* [See col. 3, lines 32-40] *by passing the syntactic possibilities (the final set of units) on to deductive token merging code. The deductive code reviews possible interpretations and determines which ones are remaining for the inductive merging processes* [See col. 5, lines 59-64].

As per claim 10, Bond teaches 'wherein generating unit associations includes identifying units that are associated with other units' *a listing of words and their associated part of speech is analyzed* [See col. 3, lines 32-40]. Bond teaches that each of word in a sentence is a '*word token*', which is equivalent to the claim limitation "unit". Bond disclosed that an article: "*United States*" [See col. 4,line 19] is tokenized into word tokens, "*United*" and "*States*". Therefore, the "*United*" is a subset of other "unit", wherein the other unit is *a "States"*.

As per claim 11, Bond teaches 'storing the unit associations to the memory' *as the syntactically marked text is output and retained in a software object to permit further analysis* [See col. 6, lines 13-15].

As per claim 12, Bond teaches 'wherein identifying associated units includes determining which units appear in queries with other units' *as a determination of which syntactic identifier is the most commonly used for a given word* [See col. 6, lines 3-5].

The italicized words from the prior art are mapped to the claim limitations.

As per claim 13, Bond teaches 'generating unit alternatives after the convergence condition has been satisfied' *as a syntactic sequence that fits the definition of a clause or consecutive tokens* [See col. 5, lines 32-35]. It should be noted that the instant specifications describes a "unit alternatives" as a "sequence of words" [See Applicant's specification page 13, line 25].

As per claim 14, Bond teaches 'wherein generating unit alternatives includes determining whether an edit distance between two units in the final set of units is smaller than a threshold value' *as the possible syntactic interpretations of the word tokens of the sentence may be deleted or replaced* by identifiers covering a smaller class of words [See col. 2, lines 7-9], 'and if so, comparing the relative frequencies of the two units' *as the series of tokens are used consecutively and compared with a first list of rules in order to produce a narrower set of possible syntactic interpretations* [See col. 4, lines 57-58, col. 2, lines 5-7].

As per claim 15, Bond teaches:

'generating unit extensions using the final set of units' *as the syntactic possibilities are passed on to deductive token merging code. The deductive code reviews possible interpretations and determines which ones are remaining for the inductive merging processes* [See col. 5, lines 59-64];

The italicized words from the prior art are mapped to the claim limitations.

'generating unit associations using the final set of units' *as a listing of words and their associated part of speech is analyzed* [See col. 3, lines 32-40] *by passing the syntactic possibilities (the final set of units) on to deductive token merging code. The deductive code reviews possible interpretations and determines which ones are remaining for the inductive merging processes* [See col. 5, lines 59-64]; and

'generating unit alternatives using the final set of units' *as a syntactic sequence that fits the definition of a clause or consecutive tokens* [See col. 5, lines 32-35]. It should be noted that the instant specifications describes a "unit alternatives" as a "sequence of words" [See Applicant's specification page 13, line 25].

As per claim 16, Bond teaches 'storing the unit extensions, the unit associations and the unit alternatives to the memory' *as the syntactically marked text is output and retained in a software object to permit further analysis* by the users [See col. 6, lines 13-17].

As per claim 17, Bond teaches:

'wherein generating unit extensions includes identifying units that are subsets of other units' *as all words in a sentence are tokenized into multiple words* [See col. 4, lines 20-21]. Because Bond teaches that each of word in a sentence is a '*word token*', which is equivalent to the claim limitation "unit". Bond disclosed that an article: "*United States*" [See col. 4,line 19] is tokenized into word tokens, "*United*" and "*States*". Therefore, the "*United*" is a subset of other "unit", wherein the other unit is *a "States",*

The italicized words from the prior art are mapped to the claim limitations.

'wherein generating unit associations includes identifying units that are associated with other units' *a listing of words and their associated part of speech is analyzed* [See col. 3, lines 32-40]. Bond teaches that each of word in a sentence is a '*word token*', which is equivalent to the claim limitation "unit". Bond disclosed that an article: "*United States*" [See col. 4,line 19] is tokenized into word tokens, "*United*" and "*States*". Therefore, the "*United*" is a subset of other "unit", wherein the other unit is *a "States",* and

'wherein generating unit alternatives includes determining whether an edit distance between two units in the final set of units is smaller than a threshold value' *as possible syntactic interpretations of the word tokens of the sentence may be deleted or replaced* by identifiers covering a smaller class of words [See col. 2, lines 7-9], 'and if so, comparing the relative frequencies of the two units' *as the series of tokens are used consecutively and compared with a first list of rules in order to produce a narrower set of possible syntactic interpretations* [See col. 4, lines 57-58, col. 2, lines 5-7].

As per claim 18, Bond teaches 'wherein validating includes for each combined unit in the second set of units' *as the processor performs matching consecutive tokens in the resulting narrower set of possible syntactic interpretations against a second set of rules through the inductive merging program* [See col. 5, lines 31-33, 51], 'comparing a frequency of occurrence of the combined unit with a frequency of occurrence of each constituent unit in the combined unit' *as the series of tokens are used consecutively and*

The italicized words from the prior art are mapped to the claim limitations.

compared with a first list of rules in order to produce a narrower set of possible syntactic interpretations [See col. 4, lines 57-58, col. 2, lines 5-7].

As per claim 19, Bond teaches 'wherein the convergence condition includes a threshold value, wherein the convergence condition is satisfied if a change in the number of units in the two second set of units between successive steps of combining and validating is smaller than or equal to the threshold value' *as possible syntactic interpretations of the word tokens of the sentence may be deleted or replaced* by identifiers covering a smaller class of words [See col. 2, lines 7-9].

As per claim 28, Steinkraus teaches the recited limitations as follows:

'a computer readable medium including code for causing a processor to generate concept units from a plurality of user search queries' *as the search queries input by the user wherein the search terms are broken into word tokens* [See col. 2, lines 40-45] and *the search engine uses the word tokens to generate a list of concept tokens* [See col. 2, line 34]. The list of content tokens is used to *search for concepts* [See col. 2, line 58-59], 'each query comprising a string of one or more words' *as the query contains words, word-like strings, numbers* [See col. 7, lines 38-39], 'wherein the code includes instructions to:

a) tokenize each query string to produce one or more tokens for each query' *as the query is read and broken into word tokens* [See abstract, lines 13-14], 'wherein said

---

The italicized words from the prior art are mapped to the claim limitations.

tokens for said queries form an initial set of units' *as the queries are broken and produced a series of word tokens* [See col. 5, lines 35-36];

'until a convergence condition is satisfied' *as the concept tokens weights are less than a threshold value* [See col. 8, lines 8-10].

Steinkraus does not explicitly teach b) combine units from the initial set of units that appear adjacent each other in a query to form a second set of units, c) validate the second set of units, d) repeat b) and c) one or more times using the second set of units in place of the initial set of units, wherein a final set of units is formed once the convergence condition has been satisfied and store the final set of units to a memory module.

However, Bond teaches, an analogous system for *tokenizing the words of the natural language* [See abstract, lines 2-3]. In particular, Bond teaches:

'b) combine units from the initial set of units that appear adjacent each other in a query to form a second set of units' *as a token merging program for merging the sequence of a phrase or clause* [See col. 5, line 51, line 35, col. 1, line 26], *the processor iterates matching consecutive tokens against the first set of rule* [See Fig. 1, # 26, col. 5, lines 26-27]. *The processor will continues to iterate through the tokenized of words until further word tokens can be made by processor* [See col. 5, lines 27-31]. Bond explained *as the processor performs matching consecutive tokens in the resulting narrower set of possible syntactic interpretations against a second set of rules* [See col. 5, lines 32-33];

The italicized words from the prior art are mapped to the claim limitations.

'c) validate the second set of units' *the processor again, processes the second set of rules until no further narrowing of possible syntactic* interpretations [See col. 5, lines 36-37];

'd) repeat b) and c) one or more times using the second set of units in place of the initial set of units' *as the process proceeds into reiteration program code* [See col. 5, lines 38-40], 'wherein a final set of units is formed once the convergence condition has been satisfied' *as the completion of the iterative inductive processing output the syntactic possible interpretations marked text* [See Fig. 1, # 38, col. 5, lines 42-44]; and

'store the final set of units to a memory module' *as the syntactically marked text is output and retained in a software object to permit further analysis by the users* [See col. 6, lines 13-17].

Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to combine the teachings of the cited references to implement the steps of combining units, validating all units that adjacent each other, and repeating the steps of combining until a convergence condition is satisfied, then storing the final set of units to memory. Bond's teachings would allow users of Steinkraus's system to readily determine the appropriate syntax for a natural language query to obtain meaningful interpretations and analyses of enormous amount of information to conveniently search for the concepts associated with any given user's needs as suggested by Bond [See col. 1, lines 29-40, col. 2, line 62-63].

As per claim 29, Bond teaches

The italicized words from the prior art are mapped to the claim limitations.

'wherein the code further includes instructions to: generate unit extensions using the final set of units' *as the completion of syntactic possibilities interpretations are passed on to deductive token merging code. The deductive code reviews possible interpretations and determines which ones are remaining for the inductive merging processes* [See col. 5, lines 59-64];

'generate unit associations using the final set of units' *as a listing of words and their associated part of speech is analyzed* [See col. 3, lines 32-40] *by passing the syntactic possibilities (the final set of units) on to deductive token merging code. The deductive code reviews possible interpretations and determines which ones are remaining for the inductive merging processes* [See col. 5, lines 59-64];

'generate unit alternatives using the final set of units' *as a syntactic sequence that fits the definition of a clause or consecutive tokens* [See col. 5, lines 32-35]. It should be noted that unit alternatives is described in the specification is a "sequence of words" [See Applicant's specification page 13, line 25]; and

'store the unit extensions, unit associations and unit alternatives to the memory module in association with the final set of units' *the syntactically marked text is output and retained in a software object to permit further analysis by the users* [See col. 6, lines 13-17].

As per claim 30, Bond teaches:

'wherein the instructions to generate unit extensions includes instructions to identify units that are subsets of other units' *as all words in a sentence are tokenized*

The italicized words from the prior art are mapped to the claim limitations.

into multiple words [See col. 4, lines 20-21]. Because Bond teaches that each of word in a sentence is a '*word token*', which is equivalent to the claim limitation "unit". Bond disclosed that an article: "*United States*" [See col. 4,line 19] is tokenized into word tokens, "*United*" and "*States*". Therefore, the "*United*" is a subset of other "unit", wherein the other unit is *a "States",*

'wherein the instructions to generate unit associations includes instructions to identify units that are associated with other units' *a listing of words and their associated part of speech is analyzed* [See col. 3, lines 32-40]. Bond teaches that each of word in a sentence is a '*word token*', which is equivalent to the claim limitation "unit". Bond disclosed that an article: "*United States*" [See col. 4,line 19] is tokenized into word tokens, "*United*" and "*States*". Therefore, the "*United*" is a subset of other "unit", wherein the other unit is *a "States",* and

'wherein the instructions to generate unit alternatives includes instructions to determine whether an edit distance between two units in the final set of units is smaller than a threshold value' *as possible syntactic interpretations of the word tokens of the sentence may be deleted or replaced* by identifiers covering a smaller class of words [See col. 2, lines 7-9], 'and if so, compare the relative frequencies of the two units' *as the series of tokens are used consecutively and compared with a first list of rules in order to produce a narrower set of possible syntactic interpretations* [See col. 4, lines 57-58, col. 2, lines 5-7].

The italicized words from the prior art are mapped to the claim limitations.

As per claim 31, Steinkraus teaches 'wherein each word comprises one or a plurality of alphanumeric characters' *as word tokens may be words, word-like strings, numbers, etc.* [See col. 7, lines 38-39].

Claims 2-5, 20-27 are rejected under 35 U.S.C. 103(a) as being unpatentable over Steinkraus (US Patent 6,363,373 B1) in view of Bond et al (US Patent 6,539,348 B1) and further in view of Whitman et al (US Patent 6,772,150 B1).

As per claim 2, Steinkraus and Bond do not teach wherein receiving includes receiving one or more query log files, each query log file including a plurality of queries. However, Whitman teaches 'receiving one or more query log files' *as a web server, maintains a daily transaction log file of all the requests it has received from web browsers* [See Fig. 3, col. 9, lines 1-3], 'each query log file including a plurality of queries' *as the daily transaction log file 135 is made up several entries, each entry containing information about a different request* [See Fig. 3,col. 9, lines 3-5]. Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made, to combine the teachings of the cited references, because Whitman's teaching of daily transaction log file that received all of the requests from web browsers would allow users of Steinkraus and Bond's system to utilize the transaction log file as a references or as a look-up table, to select the previous related search phrases to be suggested to the user when the user submits a search query which has a similar context of search terms. The related query terms suggests to the user as a part of the search refinement

---

The italicized words from the prior art are mapped to the claim limitations.

process. The goal of this process is to produce a refined search query that more narrowly specifies the user's intended requests, as explained in Whitman [See col. 1, lines 54-58, col. 2, line 51-53, col. 12, lines 49-50].

As per claim 3, Steinkraus and Bond do not teach consolidating the plurality of queries from the one or more query log files into a single consolidated query file. However, Whitman teaches 'consolidating the plurality of queries from the one or more query log files into a single consolidated query file' *as combining new data with existing data from earlier constituent time periods to form a collective search phrase table (log file)* [See Fig. 4, # 440, 450, col. 8, lines 38-40, 49-51]. Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to combine the teachings of the cited references because merging or combining of transaction log files as disclosed by Whitman's system would allow users of Steinkraus and Bond to eliminate the duplicate phrases [See col. 9, lines 48-53, col. 10, lines 45-46, col. 13, lines 30-35] so that the selection process selects related search phrases to be suggested to the user is invoked only for single term queries [See col. 12, line 64]. An important benefit of this method is that it is highly efficient, allowing the query result page to be returned without adding appreciable delay, as suggested in Whitman [See col. 12, lines 51-65].

As per claim 4, Whitman teaches 'wherein consolidating includes removing duplicates of queries and incrementing a count associated with each individual query

---

The italicized words from the prior art are mapped to the claim limitations.

each time a duplicate of said individual query is removed, wherein the consolidated file includes a list of individual queries and counts associated therewith' *as the search phrases can be done by sorting alphabetically and removing duplicate phrases* [See col.13, lines 31-35], i*f the search phrase already exists under the key term pair, the process increments the search phrase's score* [See col.10, lines 63-64].

As per claim 5, Whitman teaches 'wherein the received query log files include query log files for each day of a week' *as daily results file for 7-Feb-98* [See Fig. 6, col. 12, line 8], 'and wherein consolidating includes forming a single consolidated query file including queries for the week' *as generating a new query table for the period of the last seven days 1-Feb-98 to 7-Feb-98* [See Fig. 6, col.12, lines 11-19].

As per claim 20, Steinkraus teaches:

'receiving an individual query from a user' *as a search query input by the user* [See Fig. 3, col. 7, line 36];

'identifying one or more units in the individual query' *as the query is read and broken into word tokens* [See abstract, lines 13-14]; and

'unit alternatives' *as a syntactic sequence that fits the definition of a clause or consecutive tokens* [See col. 5, lines 32-35]. It should be noted that the instant specifications describes a "unit alternatives" as a "sequence of words" [See Applicant's specification page 13, line 25].

The italicized words from the prior art are mapped to the claim limitations.

Steinkraus does not explicitly teach using one or more of the unit extensions, unit associations stored in the memory in association with the one or more units identified in the individual query. However, Bond teaches 'using one or more of the unit extensions' *as the syntactic possibilities are passed on to deductive token merging code. The deductive code reviews possible interpretations and determines which ones are remaining for the inductive merging processes* [See col. 5, lines 59-64], 'unit associations' *as a listing of words and their associated part of speech is analyzed* [See col. 3, lines 32-40] *by passing the syntactic possibilities (the final set of units) on to deductive token merging code. The deductive code reviews possible interpretations and determines which ones are remaining for the inductive merging processes* [See col. 5, lines 59-64] and 'stored in the memory in association with the one or more units identified in the individual query' *as the syntactically marked text is output and retained in a software object to permit further analysis* [See col. 6, lines 13-15]. Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to combine the teachings of the Steinkraus and Bond's system, because the beneficial of Bond's teaching of using the unit extension, unit association stored in the memory in association with units identified in the individual query would offer users of Steinkraus's system to directly access to the syntax parse independent without the need for customizing of any semantic information such as happens in noun and verb classes, as explained in Bond [See col. 6, lines 16-21].

Steinkraus and Bond do not explicitly teach determining one or more suggestions to provide to the user responsive to the query. However, Whitman teaches 'determining

---

The italicized words from the prior art are mapped to the claim limitations.

one or more suggestions to provide to the user responsive to the query' *as when a user submits a search query, a query processing component uses the search phrases table 137 of Figure 1, to look up one or more related search phrases to suggest to the user as alternative queries* [See col. 2, lines 52-54]. Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made, to combine the teachings of the cited references because Whitman's system would allow users of Steinkraus and Bond's system to utilize a search phrases table 137 to select the related search phrases to be suggested to the user when the user submits a search query which has a similar context of search terms. The related query terms suggests to the user as a part of the search refinement process. The goal of this process is to produce a refined search query that more narrowly specifies the user's intended requests, as explained in Whitman [See col. 1, lines 54-58, col. 2, line 51-53, col. 12, lines 49-50].

As per claim 21, Steinkraus discloses the recited limitations as follows:

'a system for generating concept units from user search queries' *as the search queries input by the user wherein the search terms are broken into word tokens* [See col. 2, lines 40-45] and *the search engine uses the word tokens to generate a list of concept tokens* [See col. 2, line 34]. The list of content tokens is used to *search for concepts* [See col. 2, line 58-59], the system comprising:

'a memory unit' *as a memory includes RAM and ROM* [See Fig. 1, # 22, # 24 and # 25]; and

---

The italicized words from the prior art are mapped to the claim limitations.

'a processing module' *as the program modules* [See col. 4, line 19], 'configured to receive one or more query files' *as preprocessing search queries input by the users* [See col. 7, line 30 and line 36], 'each query including a string of one or more words' *as the query contains words, word-like strings, numbers* [See col. 7, lines 38-39], 'and wherein the processing module is further configured to:

tokenize each query from the query files' *as the query is read and broken into word tokens* [See abstract, lines 13-14] 'to produce an initial set of units' *as the queries are broken and produced a series of word tokens* [See col. 5, lines 35-36]; and

'thereafter, iteratively, until a convergence condition is satisfied' *as the concept tokens weights are less than a threshold value* [See col. 8, lines 8-10].

Steinkraus does not explicitly teach combine units from the initial set of units that appear adjacent each other in a query to form a second set of units, validate the second set of units, wherein the second set of units is used for each iteration, and once the convergence condition has been satisfied, store a final set of units to the memory unit.

However, Bond teaches, an analogous system for *tokenizing words of a natural language* [See abstract, lines 2-3]. In particular, Bond teaches:

'combine units from the initial set of units that appear adjacent each other in a query to form a second set of units' *as a token merging program for merging the sequence of a phrase or clause* [See col. 5, line 51, line 35, col. 1, line 26], *the processor iterates matching consecutive tokens against the first set of rule* [See Fig. 1, # 26, col. 5, lines 26-27]. *The processor will continues to iterate through the tokenized of words until further word tokens can be made by processor* [See col. 5, lines 27-31].

The italicized words from the prior art are mapped to the claim limitations.

Bond explained *as the processor performs matching consecutive tokens in the resulting narrower set of possible syntactic interpretations against a second set of rules* [See col. 5, lines 32-33]; and

'validate the second set of units' *as the processor again, processes the second set of rules until no further narrowing of possible syntactic* interpretations [See col. 5, lines 36-37], 'wherein the second set of units is used for each iteration' *as the process proceeds into reiteration program code* [See col. 5, lines 38-40]; and

'once the convergence condition has been satisfied' *as the completion of the iterative inductive processing output the syntactic possible interpretations marked text* [See Fig. 1, # 38, col. 5, lines 42-44], 'store a final set of units to the memory unit' *as the syntactically marked text is output and retained in a software object to permit further analysis by the users* [See col. 6, lines 13-17].

Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to combine the teachings of the cited references to implement the steps of combining units, validating all units that adjacent each other, and repeating the steps of combining until a convergence condition is satisfied, then storing the final set of units to memory. Because Bond's teachings would allow users of Steinkraus's system to readily determine the appropriate syntax for a natural language query to obtain meaningful interpretations and analyses of enormous amount of information to conveniently search for the concepts associated with any given user's needs as suggested by Bond [See col. 1, lines 29-40, col. 2, line 62-63].

The italicized words from the prior art are mapped to the claim limitations.

Steinkraus and Bond do not teach receive one or more query log files, each query log file including a plurality of queries. However, Whitman teaches 'receive one or more query log files' *as a web server, maintains a daily transaction log file of al the requests it has receives from web browsers* [See Fig. 3, col. 9, lines 1-3], 'each query log file including a plurality of queries' *as the daily transaction log file 135 is made up several entries, each entry containing information about a different request* [See Fig. 3,col. 9, lines 3-5]. Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made, to combine the teachings of the cited references, because Whitman's teaching of daily transaction log file that received all of the requests from web browsers would allow users of Steinkraus and Bond's system to utilize the transaction log file as a references or as a look-up table, to select the previous related search phrases to be suggested to the user when the user submits a search query which has a similar context of search terms. The related query terms suggests to the user as a part of the search refinement process. The goal of this process is to produce a refined search query that more narrowly specifies the user's intended requests, as explained in Whitman [See col. 1, lines 54-58, col. 2, line 51-53, col. 12, lines 49-50].

As per claim 22, Steinkraus teaches 'one or more query log file sources for providing the query log file' *as a distributed computing environments* [See col. 3, lines 53-55].

As per claim 23, Bond teaches

The italicized words from the prior art are mapped to the claim limitations.

'wherein the processing module is further configured to: generate unit extensions using the final set of units' *as the completion of syntactic possibilities interpretations are passed on to deductive token merging code. The deductive code reviews possible interpretations and determines which ones are remaining for the inductive merging processes* [See col. 5, lines 59-64];

'generate unit associations using the final set of units' *as a listing of words and their associated part of speech is analyzed* [See col. 3, lines 32-40] *by passing the syntactic possibilities (the final set of units) on to deductive token merging code. The deductive code reviews possible interpretations and determines which ones are remaining for the inductive merging processes* [See col. 5, lines 59-64];

'generate unit alternatives using the final set of units' *as a syntactic sequence that fits the definition of a clause or consecutive tokens* [See col. 5, lines 32-35]. It should be noted that unit alternatives is described in the specification is a "sequence of words" [See Applicant's specification page 13, line 25]; and

'store the unit extensions, unit associations and unit alternatives to the memory in association with the final set of units' *the syntactically marked text is output and retained in a software object to permit further analysis by the users* [See col. 6, lines 13-17].

As per claim 24, Steinkraus and Bond do not teach wherein the received query log files include query log files for each day of a week, and wherein the processing module is further configured to consolidate the query log files into a single consolidated query file consisting of queries for the week. However, Whitman teaches 'wherein the

---

The italicized words from the prior art are mapped to the claim limitations.

received query log files include query log files for each day of a week' *as daily results file for 7-Feb-98* [See Fig. 6, col. 12, line 8], 'and wherein the processing module is further configured to consolidate the query log files into a single consolidated query file consisting of queries for the week' *as generating a new query table for the period of the last seven days 1-Feb-98 to 7-Feb-98* [See Fig. 6, col.12, lines 11-19]. Thus, it would have been obvious to one of ordinary skill in the art at the time the invention was made to combine the teachings of the cited references because merging or combining of transaction log files to produce a new search phrase table as disclosed by Whitman's system [See col. 10, lines50-51] is equivalent to the claim "consolidate the query log files into a single consolidated query file", would allow users of Steinkraus and Bond to eliminate the duplicate phrases [See col. 9, lines 48-53, col. 10, lines 45-46, col. 13, lines 30-35] so that the selection process selects related search phrases to be suggested to the user is invoked only for single term queries [See col. 12, line 64]. An important benefit of this method is that it is highly efficient, allowing the query result page to be returned without adding appreciable delay, as suggested in Whitman [See col. 12, lines 51-65].

As per claim 25, Whitman teaches 'wherein the processing module consolidates by removing duplicates of queries and incrementing a count associated with each individual query each time a duplicate of said individual query is removed, wherein the consolidated file includes a list of individual queries and counts associated therewith' *as the search phrases can be done by sorting alphabetically and removing duplicate*

The italicized words from the prior art are mapped to the claim limitations.

phrases [See col.13, lines 31-35], i*f the search phrase already exists under the key term pair, the process increments the search phrase's score* [See col.10, lines 63-64].

As per claim 26, Whitman teaches 'wherein the processing module determines a frequency of occurrence for each unit using the counts associated with the queries' *as a score is maintained for each search phrase to reflect how many times the search phrase appeared in the transaction log* [See col. 11, lines 4-6], 'and wherein the processing modules stores the unit frequencies to the memory unit in association with the final set of units' *as the final values of the scores taken over total number of days and are stored within the search phrase table* [See col. 7, lines 6-7].

As per claim 27, Steinkraus teaches 'wherein the memory unit and processing module are implemented in a search server device in a network' *as the computing distributed environments where tasks are performed by remote processing devices that are linked through a communications network* [See col. 3, lines 53-55].

Conclusion

The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

U.S Patent 6,006,221, issued to Liddy et al on Dec. 21, 1999. The subject matter disclosed therein in is a user submits a natural language query. The system processes

---

The italicized words from the prior art are mapped to the claim limitations.

the query to generate a language-independent conceptual representation of the subject content of the query and display the reformulate queries output results to the user.

U.S Patent Application No. 2003/0069880 A1, Publication date April 10, 2003, to Harrison et al. The subject matter disclosed therein is tokenizing an incoming query and compared the tokenized representation against a collection of query templates. The most relevant results to the incoming query returns to the user.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to DEBBIE M LE whose telephone number is 703-308-6409. The examiner can normally be reached on 8:30-5:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, JOHN BREENE can be reached on 703-305-9790. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR.

The italicized words from the prior art are mapped to the claim limitations.

Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

DEBBIE M LE
Examiner
Art Unit 2177

Debbie Le

Aug. 20, 2004.

The italicized words from the prior art are mapped to the claim limitations.